The Impact of Lightning Data Assimilation on a Winter Storm Simulation over the North Pacific Ocean

ANTTI T. PESSI AND STEVEN BUSINGER
University of Hawaii at Manoa, Honolulu, Hawaii

(Manuscript received 26 August 2008, in final form 26 January 2009)

ABSTRACT
In this paper, the potential of lightning data assimilation to improve NWP forecasts over data-sparse oceans is investigated using, for the first time, a continuous, calibrated lightning data stream. The lightning data employed in this study are from the Pacific Lightning Detection Network/Long-Range Lightning Detection Network (PacNet/LLDN), which has been calibrated for detection efficiency and location accuracy. The method utilizes an empirical lightning–convective rainfall relationship, derived specifically from North Pacific winter storms observed by PacNet/LLDN. The assimilation method nudges the model’s latent heating rates according to rainfall estimates derived from PacNet/LLDN lightning observations. The experiment was designed to be employed in an operational setting. To illustrate the promise of the approach, lightning data from a notable extratropical storm that occurred over the northeast Pacific Ocean in late December 2002 were assimilated into the fifth-generation Pennsylvania State University–NCAR Mesoscale Model (MM5). The storm exhibited a very electrically active cold front with most of the lightning observed 300–1200 km away from the storm center. The storm deepened rapidly (12 hPa in 12 h) and was poorly forecast by the operational models. The assimilation of lightning data generally improved the pressure and wind forecasts, as the validation of the model results using available surface and satellite data revealed. An analysis is presented to illustrate the impact of assimilation of frontal lightning on the storm development and dynamics. The links among deep convection, thermal wind along the front, and cyclogenesis are explicitly explored.

1. Introduction
Eastward-propagating extratropical storms over the North Pacific Ocean can bring hazardous weather conditions to the West Coast of North America. These storms are often poorly forecast because the lack of observations over the Pacific Ocean results in poorly described initial conditions in NWP models (McMurdie and Mass 2004). In turn, inadequate initial conditions along the West Coast can lead to error propagation in model forecasts downstream over the mainland United States. A notable example was a poorly forecast extratropical cyclone that deepened rapidly as it approached the West Coast in December 2002. The National Centers for Environmental Prediction (NCEP) Global Forecast System (GFS) produced a 10-hPa error in the 12-h central pressure forecast for this storm. An interesting characteristic of this system was a band of active thunderstorms along the storm’s cold front.
Diabatic heating sources, especially latent heat release in deep convective clouds, can play an important role in storm development and dynamics (e.g., Anthes et al. 1983; Brennan and Lackmann 2005). The inadequate initial conditions in the moisture and vertical motion fields of the NWP models result in the well-documented spinup problem (e.g., Davidson and Puri 1992). This can subsequently result in large errors in storm central pressure and rainfall forecasts (e.g., McMurdie and Mass 2004). Specifying diabatic heating sources or moisture distribution in the early hours of a forecast can improve the model’s performance (Bauman et al. 1997; Businger et al. 2005).
Lightning climatology, as observed by the satellite-borne Lightning Imaging Sensor (LIS) and Optical Transient Detector (OTD), shows a winter maximum along the storm track extending from Japan to the west coast of North America (Christian et al. 2003). Recently, data from the Pacific Lightning Detection Network/Long-Range Lightning Detection Network (PacNet/
LLDN; hereafter just PacNet) have supported this observation (Pessi et al. 2009; Pessi and Businger 2009). Lightning observations over the North Pacific Ocean offer important additional information about convective activity over the data-sparse ocean. Unfortunately, LIS’s observations only cover the latitudes between \(~38^\circ S\) and \(~38^\circ N\), with a few minutes of observation time per day over each point. In contrast, PacNet offers a continuous, calibrated lightning data stream over the North Pacific Ocean that is suitable for lightning data assimilation. Some low-orbit satellites are equipped with microwave imagers and precipitation radars and offer information concerning convective activity, but they typically have only twice-per-day coverage. Geostationary satellites offer continuous monitoring of clouds and water vapor on visible and infrared channels, but cirrus anvils often obscure convective activity.

Several previous studies have investigated the impact of assimilation of remotely sensed lightning, satellite, and radar data using Newtonian nudging and other methods. Manobianco et al. (1994) assimilated satellite-derived precipitation into a regional-scale numerical model. The assimilation was implemented by scaling the model-generated profiles of latent heating according to rainfall rates derived from satellite observations. At grid points where the model was not producing precipitation, the vertical distribution of latent heating was specified using model-generated profiles at adjacent grid points with similar rainfall rates as estimated from satellite observations.

Alternatively, observed rainfall rates can be assimilated by “inverting” or “reversing” the precipitation parameterization. The scheme adjusts the moisture (and to some extent the temperature) profiles so that the convective parameterization scheme (CPS) reproduces the observed rainfall rate. This type of scheme has been used to initialize the model to reduce the spinup time (e.g., Donner 1988; Puri and Miller 1990).

Jones and Macpherson (1997a) used latent heat nudging to assimilate surface rain rate estimates derived from radar data over the Met Office (UKMO) mesoscale model. The model-generated profiles of latent heating were scaled by the ratio of observed to model precipitation rates. The scheme increased the forecast skill for precipitation distribution in the first 6–9 h of the forecast.

Several other studies have shown the skill of the nudging method, where rain rates from various sources were assimilated into NWP models to improve forecasts of various storm systems—including tropical cyclones (e.g., Karyampudi et al. 1998; Davolio and Buzzi 2004).

Jones and Macpherson (1997b) assimilated long-range lightning data from the UKMO Arrival Time Difference (ATD) network into a limited-area model. Their method utilized a latent heat nudging scheme that was originally developed to assimilate precipitation observations derived from the U.K. weather radar network (see earlier reference to Jones and Macpherson 1997a). They used the correlation between flash rate and precipitation rate \(R = 4.3F\) found by Buechler et al. (1994), where \(R\) is rain rate (mm h\(^{-1}\)) and \(F\) is flash rate per 15-min interval per 10-km square. They did not see any consistent improvement in forecast skill and suggested that this may have been due to the uncertainties in both the detection efficiency of the ATD network and the lightning–rainfall relationship.

Alexander et al. (1999) investigated the impact of assimilation of precipitation data from various sources on simulations of the 1993 Superstorm. They used lightning data from the National Lightning Detection Network (NLDN) and UKMO VLF networks and rainfall data from the Special Sensor Microwave Imager (SSM/I) and Geostationary Operational Environmental Satellite infrared sensor (GOES-IR). They derived a lightning–rainfall relationship using the available SSM/I and lightning data over the storm. They found that the assimilation of satellite-derived rainfall rates alone was not sufficient to reproduce the observed development of the storm; the cyclogenesis occurred at a time inconveniently placed between SSM/I overpasses (\(~12\)-h intervals), and GOES-IR imagery (3-h intervals) only responded to the convection 6 h after it actually occurred, when tall clouds and cirrus anvils had formed. However, the assimilation of continuous long-range lightning data resulted in dramatic improvement of the model prediction.

Chang et al. (2001) used sferics observations from a VLF lightning detection network and rainfall data from the Tropical Rainfall Measuring Mission (TRMM) satellite to derive the relationship between convective rainfall and lightning for the 1998 Groundhog Day storm. Assimilation of latent heating profiles derived from the lightning data, along with SSM/I and TRMM precipitation and integrated water vapor data, resulted in improved short-term (9–18 h) forecasts.

Papadopoulos et al. (2005) assimilated lightning data from a long-range Zeus network operating in Europe. They nudged model-generated moisture profiles to empirical profiles according to lightning intensity. The empirical humidity profiles were produced from atmospheric soundings during thunderstorm days. The assimilation method significantly improved convective rainfall prediction during the assimilation period and maintained the improvement in short-term forecasts up to 12 h.

Pessi et al. (2006) used the MM5’s (defined in section 2b) four-dimensional data assimilation module to nudge the model’s vertical moisture profiles according to the observed lightning rate at each grid point. This method also...
used a lightning–convective rainfall relationship, but an extra step was required to convert the rainfall rates to vertical mixing ratio profiles.

Mansell et al. (2007) used lightning data from NLDN and Lightning Mapping Array in the U.S. Midwest. They used another approach and added water vapor to the boundary layer in lightning-active areas. Water vapor was added in small increments up to 1 g kg$^{-1}$ until cloud depth reached 7 km and vertical velocity was 10 m s$^{-1}$. Thus, the convective parameterization scheme was forced to produce convection in the lightning active areas. Conversely, the CPS can be prevented from producing vigorous convection where lightning is not observed using several options. The updraft width can be reduced, increasing the entrainment and dilution of updraft air or by limiting the vertical velocity given to parcels and thus making it harder to reach the LFC. Another option is to skip the CPS altogether in that column.

Earlier studies have found that assimilating latent heat in the correct location is more important than accurate estimation of the precipitation intensity (Manobianco et al. 1994; Chang et al. 2001). These results are encouraging for lightning data assimilation because they may relax the requirements for quantifying lightning rates and deriving the lightning–rainfall relationship. In this study, the potential of lightning data assimilation to improve NWP forecasts over data-sparse oceans is investigated using several options. The updraft width can be reduced, increasing the entrainment and dilution of updraft air or by limiting the vertical velocity given to parcels and thus making it harder to reach the LFC. Another option is to skip the CPS altogether in that column.

The results shown in Fig. 1 were adjusted to take into account the smaller model grid size used in this study (0.25°) as compared to grid size used to derive the curve (0.5°). A uniform lightning density over the grid cell is assumed; thus, the lightning rates are divided by four. Finally, the convective rainfall rates derived from lightning observations are written to an assimilation file, covering each grid point and time step during the assimilation period. The data at every 81-s time step consider lightning observations occurring ±15 min of the time step. Before the model run is started, the model initialization process reads the assimilation file into internal arrays.

b. Model setup and implementation of the lightning data assimilation method

The model used in this study is the fifth-generation Pennsylvania State University–National Center for Atmospheric Research (NCAR) Mesoscale Model (MM5), version 3 (Grell et al. 1994). The horizontal resolution of the model is 27 km and it has 39 vertical levels. The model is nonhydrostatic and the time step used was 81 s. A nesting option, available for the model, was not used to simplify physical interpretation of the results. The Kain–Fritsch (KF) convective parameterization scheme was used (Kain and Fritsch 1993). Boundary conditions were provided by the Global Forecast System model and were read every 6 h. The spectral GFS model adopts a
triangular truncation of 254 waves in the horizontal with 64 vertical levels (at the time of the case study forecasts). The corresponding horizontal resolution is \( \approx 55 \text{ km or } 0.5^\circ \).

The lightning data assimilation (LDA) method was implemented in the KF convective parameterization scheme by modifying the model-generated vertical latent heating profiles. The method scales the model’s vertical latent heating profiles at each grid point and model level depending on the ratio of rainfall predicted by the model to rainfall derived from lightning data. The scheme was successfully used by Manobianco et al. (1994) to assimilate satellite-derived rainfall rates into a mesoscale model during a rapidly intensifying extratropical cyclone. A similar scheme was later used by Karyampudi et al. (1998), Alexander et al. (1999), and Chang et al. (2001) to assimilate rainfall rates derived from satellite and/or lightning data.

The Manobianco et al. (1994) scheme considers three regimes. (i) If both the rainfall \( R_o \) derived from lightning rates and model-predicted rainfall \( R_m \) are greater than zero \( (R_o > 0, R_m > 0) \), then the scaling coefficient \( c = (R_o - R_m)/R_m \). The adjusted latent heating at each model level \( i \) is \( \Delta T_i = (1 + c)\Delta T_i^{\text{mdl}} \). To prevent excessive latent heating values and model instability, the scaling coefficient was limited to 3. Scaling is done only if the lightning-estimated rain rate is greater than the
model-produced rain rate. This approach is different than in Manobianco et al. (1994) and was taken because it is possible to have significant warm rainfall with low lightning rates (or absence of lightning). Moreover, in areas with observed lightning, the model rainfall was consistently less than that suggested by the lightning–rainfall rate relationship. (ii) If the observed lightning rate was zero ($R_o = 0$), no assimilation was done, as the absence of lightning does not imply the absence of rain. (iii) If the rain rate estimated from lightning observations at any grid point was greater than zero, but the model rain rate was zero ($R_o > 0$, $R_m = 0$), a search algorithm was used. Initially the algorithm searches adjacent model grid points for rainfall rates ($\pm 1$ mm h$^{-1}$) similar to those estimated. First, the grid points within a 10-grid length radius are searched and farther grid points are gradually included in the search until a match is found. If no match is found, model rain rates within $\pm 2$, 3, 5, 10, 15, and 20 mm h$^{-1}$ of those estimated are gradually included in the search. After the match is found, the vertical latent heating profile from the matching grid point is used. Because latent heat is insensitive to changes in temperature and the saturation mixing ratio is strongly a nonlinear function of temperature, the moisture profile from the matching grid point is not used because it can originate from a different thermodynamic environment. Instead, the levels where the heating rate is positive are assumed to be saturated. The assimilation method does not adjust the model rain rate per se, but it allows the KF scheme to generate
rainfall rates depending on the original conditions and the adjusted latent heat rates and moisture in the third regime.

A radius-of-influence value was set equal to the model grid length (\(0.25\)). If a flash is observed within the radius of influence from a grid point, it has an impact on that particular grid point. This approach was taken to spread the influence of lightning observations beyond a single grid point in cases of sporadic oceanic lightning.

In the operational system at the University of Hawaii, the MM5 model run starts \(~8\) h after the analysis time (Cherubini et al. 2006). MM5 uses the initial conditions from the GFS model, which are delivered \(~3.5\) h after the analysis time. After the initial MM5 analysis is completed, the time integration begins. Lightning data are disseminated in near-real time, thus allowing assimilation of data during the initial 7–\(8\) h of the forecast period. Lightning data were assimilated for the first 8 h of the model simulation. This time period was chosen to replicate conditions in a real-time operational forecast setting. In other words, the results shown in this study could have been produced operationally if the assimilation system had been implemented at the time.

The method described here is a four-dimensional data assimilation technique that nudges the model’s tendency equations at run time. The method imposes the effects of convection but is not a Newtonian-type nudging (or relaxation) method, which uses a tunable nudging coefficient that represents the relaxation time scale. The advantages of long-range lightning data are that they are continuous and available in near-real time, making them suitable for the approach presented in this paper. Because of the low computational cost of creating the assimilation input file, the cutoff time of the lightning data that are assimilated can be set to just minutes before the model run is started in an operational setting. Therefore, the cutoff time for the lightning data in the LDA method can be set much later than the cutoff time for other (e.g., satellite) data using “conventional” assimilation systems, such as three-dimensional variational data assimilation (3DVAR) or the Local Analysis and Prediction System (LAPS). Moreover, the LDA method can be used with any assimilation system. Because both the control and LDA forecasts use the same initial and boundary conditions, the LDA method demonstrates the value of lightning observations as an additional data source in an operational setting.

c. Verification of the results

The results were verified using available surface and satellite data. The surface analyses in Fig. 2 are based on the original analysis charts issued by the National Weather Service Honolulu Forecast Office at \(6\)-h intervals. However, for legibility only a subset of the available observations is plotted.

Together with the surface observations, satellite data from SeaWinds on the Quick Scatterometer (QuikSCAT), the Defense Meteorological Satellite Program (DMSP) SSM/I, and the Advanced Microwave Scanning Radiometer for Earth Observing System (AMSR-E) instruments were used to validate the modeled wind fields and to locate the center of the storm. Wind speed data are available from SSM/I (Goodberlet et al. 1989) and AMSR-E (Shibata 2006), whereas QuikSCAT provides both wind speed and direction data (Jet Propulsion Laboratory 2001). However, the quality of the wind direction measurements depends on the wind speed. The accuracy of the wind direction derived from QuikSCAT at low wind speeds (\(~5\) m s\(^{-1}\)) is worse than that at moderate to high wind speeds (Ebuchi et al. 2002). Therefore, the location of the wind speed minimum at each overpass was assumed to collocate approximately with the storm center. This assumption gives reasonable estimates of the storm-center location over the data-sparse ocean regions (e.g., Chelton et al. 2006). The rain contamination flags were ignored because the purpose of the validation is not to compare individual wind barbs but to show the large-scale wind field. The spatial resolution of all the wind data were \(~25\) km, matching well with the model resolution of \(27\) km.

In addition, hourly sea level pressure and wind data from the National Data Buoy Center were used for model validation. Buoy 46005, positioned at 46.05°N, 131.02°W, was located \(~300\) km to the east of the storm center at 1200 UTC 19 December 2002.

3. Results

Beginning on 18 December 2002, an extratropical cyclone started to produce lightning as it moved across the eastern Pacific Ocean. The surface analysis valid at 0000 UTC 18 December shows a trough located at 46°N, 146°W with sea level pressure of 994 hPa (Fig. 2a). This is approximately 1700 km west of the U.S. West Coast and \(~3000\) km north-northeast of Hawaii. Only a few lightning strikes were observed at this time (Fig. 2b). Six hours later, the storm had deepened to 990 hPa (center at 47°N, 144°W) and increasing lightning activity was observed along the cold front. By 1200 UTC 18 December, the storm had developed a well-defined closed circulation with analyzed central pressure of 986 hPa (Fig. 2c), and abundant lightning was observed along the cold front (Fig. 2d). At 1800 UTC 18 December, the central pressure was 985 hPa and a secondary low center (992 hPa) had developed to the south-southeast of the original center (not shown). The storm continued to move slowly.
eastward, and 6 h later, at 0000 UTC 19 December, the
two centers merged, with the secondary center becoming
the primary center (Fig. 2e). A storm central pressure of
984 hPa was analyzed with continued prolific lightning
along the cold front (Fig. 2f). The low center slowed and
then nearly stalled near 45°N, 135°W during the following
12 h (Fig. 2g), while the cold front progressed eastward
and continued to produce significant lightning rates. In
addition, enhanced lightning was observed in the unstable
cold air west of the cold front (Fig. 2h).

The history of the analyzed storm central pressure is
based on the best subjective analysis of the surface data
(e.g., Fig. 2) and may have some uncertainty. Therefore,
the main purpose of this presentation is to show the
impact of frontal lightning on the storm development
and the general trend of the sea level pressure. That said,
the satellite-based wind observations support the va-

dility of the surface analyses, as will be discussed next
and in section 3b.

**a. 24-h model run**

The storm deepened rapidly (from 984 to 972 hPa)
between 0000 and 1200 UTC 19 December 2002, as
determined from the surface analyses. A 12-h MM5
control forecast initialized at 0000 UTC and valid at
1200 UTC 19 December 2002 showed 982-hPa storm
central pressure (Fig. 3), whereas the analyzed value was
10 hPa lower (Figs. 2g, 3). The NCEP GFS Aviation
Model (GFS/AVN) and Eta models showed similar er-

FIG. 3. Comparison of analyzed storm central pressure (solid)
with that predicted by the MM5 (control; dashed) and LDA
(dotted) runs. The model was initialized at 0000 UTC 19 Dec
2002 and run for 24 h. The observed values are based on the NWS
North Pacific surface analyses every 6 h. The arrow shows the LDA
period.

ers. The GFS (Eta) 12-h forecast was 10 (9) hPa too high.
When lightning data were assimilated into MM5, the 12-h central pressure forecast error reduced to 5 hPa (Fig. 3). Three hours later (15-h forecast), the MM5 central pressure forecast with LDA reached 975 hPa. The LDA run simulates the rate of the rapid pressure drop relatively well between 0000 and 1200 UTC but starts the deepening 3 h too late, resulting in too-high pressure values. It also raises the pressure more slowly than analyzed (Fig. 3). In contrast, the control run fails to simulate the deepening of the storm and keeps the pressure between 984 and 981 hPa.

The storm location and central pressure were relatively well initialized in the model, compared to the analysis (Fig. 4). During the first 6 h, both the control and LDA forecasts place the storm close to the observed location. Thereafter, the LDA run moves the storm to the northeast, whereas in the control run the storm stays relatively stationary until 1200 UTC and moves to the north thereafter. Both model runs place the storm too far to the north and make a loop to the west. This feature was not seen in the satellite and surface data (Fig. 4). The more northeasterly location of the storm in the LDA run between 0900 and 1200 UTC is a result of the latent heating effects, which will be discussed later in this section.

The sea level pressures at 0000 UTC 20 December of the control and LDA runs were 983 and 975 hPa, respectively (Fig. 3). A more detailed analysis, based on data from a QuikSCAT overpass at 0310 UTC 20 December, reveals that the lower pressure resulted in an enhanced wind field in the LDA run that matches the QuikSCAT wind field better than the control forecast (Fig. 5). Specifically, the area to the south and southwest is better simulated by the LDA run, with the wind speed matching the QuikSCAT observations relatively well. The band of higher winds on the east side of the storm is better simulated with LDA, although the predicted wind speeds are still slightly too low. In contrast, the control run underforecasts the wind speeds over practically the entire domain shown in Fig. 5a, with the largest errors appearing generally on the north side of the storm.

In comparison with observations from the nearest buoy ~300 km east of the storm center, the control run
simulated the deepening of the storm well until 1200 UTC, whereas the sea level pressure from the LDA forecast was 5 hPa too low (Fig. 6a). Between hours 15 and 24, the pressure forecast from the LDA run performed better than the control forecast, with errors of 0–2 and 4–5 hPa, respectively. Wind data from the same buoy location showed that the LDA and control runs were relatively similar (within 2 m s$^{-1}$) between forecast hours 0 and 12. Consistent with pressure data, the LDA forecast between hours 15 and 24 tracked closer to observed than did the control forecast (Fig. 6b).

The biggest differences in rainfall between the 12-h LDA and control runs were found near the cold front and over the cold pool, which were very lightning-active, but also around the storm center to the north and west (Fig. 7). It is interesting to note that very little lightning was observed north of 45$^\circ$N during the assimilation period (Fig. 2f). However, the LDA run shows a large increase in rainfall relative to control run in some areas between 45$^\circ$ and 50$^\circ$N. It is suggested that this moisture was advected to these areas from the lightning active regions to the south, as described below. Note that in some areas in Fig. 7 the increase in rainfall occurs adjacent to areas with decreased rainfall. This likely indicates a shift in the rainfall patterns rather than just an increase in rainfall amount. The largest differences in sea level pressure between the 12-h LDA and control runs are near the center of the LDA-predicted low and over the cold front to the southeast of the center (Fig. 7c).

If the majority of the lightning was observed over a narrow cold front hundreds of kilometers (~300–1200 km) away from the storm center, what caused the storm central pressure to drop? Figures 8a–d show the evolution of the vertically integrated virtual temperature ($T_v$) field between 0000 and 1200 UTC 19 December. Virtual temperature was chosen to illustrate the evolution of warmer and moister air mass resulting from the assimilation of lightning data. The warming, as diagnosed by $T_v$, impacts the sea level pressure because the hydrostatic part of the atmospheric pressure at any given point is equal to the weight of the column of air overlying the point (vertically integrated density), and warmer air is less dense for a given pressure. The 3-h forecast shows increased $T_v$ values over the cold front and to the east of the storm center. Six hours after the start of the run, $T_v$ has further increased over the cold front and cold pool area, and some increase can be also seen near the storm center. At 0900 UTC, the lightning data assimilation has stopped, but high-$T_v$ air has advected over the storm center, resulting in slightly reduced sea level pressure. At 1200 UTC, high-$T_v$ air lies over the storm center from the surface up to ~300 hPa, with the largest anomalies between 400 and 700 hPa (Fig. 9).

A 24-h time series in Fig. 10a shows the number of lightning strikes over the area and the difference of the vertically integrated $T_v$ between the LDA and control forecasts in the storm center. During the first 6 h, while LDA was still running, the $T_v$ difference was very small over the storm center. After 6 h, $T_v$ started to rise when the air from the cold front that had been heated by the LDA started to advect over the storm center (as in Fig. 8). Simultaneously, the 950–500-hPa thickness started to increase and the sea level pressure began to fall (Fig. 10b).

Using a form of hypsometric equation (Wallace and Hobbs 1977), we can estimate the impact of the change in virtual temperature and 500-hPa height, resulting from the lightning data assimilation, on the storm central pressure (Businger 1987). We can write

$$p_{slv} = p_z \exp \left( \frac{g_0 Z}{R_d \bar{T}_v} \right),$$

where $p_{slv}$ is the sea level pressure, $p_z$ is the pressure at level $z$, $g_0$ is a gravitational constant, $Z$ is the height of the $p_z$ pressure surface, $R_d$ is a gas constant for dry air, and $\bar{T}$ is the average virtual temperature between sea level and $p_z$. For 500-hPa height, we use modeled values.
of 5247 and 5267 m for the LDA and control runs, respectively. Using an average $T_y$ value of 267.5 K from the LDA model sounding, and after substitution, we see that a $T_y$ difference of 1.2 K (from Fig. 10) and $Z_{500}$ difference of 20 m at 1200 UTC correspond to sea level pressure difference of $\sim$5 hPa, resulting in sea level pressure values of 977 and 982 hPa for the LDA and control runs, respectively (Fig. 3).

The conclusion of this analysis is that the assimilation method resulted in an adjustment of the mass (and to some extent the moisture) field in the lightning-active areas. The model responded to the altered mass

![Figure 7](image-url)
field by geostrophically adjusting the pressure and wind fields. The assimilation increased the temperature gradient across the front, with thermal wind balance ensuring an increase in the along-front winds (Fig. 11). This enhanced the advection of warm air over the storm center and reduced the sea level pressure hydrostatically.

b. 36-h model run

The impact of LDA on a longer simulation was investigated by starting the model run 12 h earlier, at 1200 UTC 18 December 2002, and similarly assimilating lightning data during the first 8 h of the model run. It should be noted that both the control and LDA forecasts
used the initial conditions from the 12-h MM5 control forecast from the previous cycle (0000 UTC) and boundary conditions from the 0000 UTC GFS forecast because data from the 1200 UTC GFS run were unavailable to create initial and boundary conditions. This approach introduces more uncertainty into both control and LDA forecasts but nevertheless allows the impact of the LDA method to be further investigated.

Frequent lightning was observed over the cold front of the storm at the time of the model run start (Fig. 2d). The front moved east and north and remained very lightning-active throughout the LDA period. The MM5 24-h forecast with LDA shows increased rainfall rates over the cold pool and around the storm center compared to the control run, and the cold front has shifted ~2° to the east when compared to the control run (Fig. 12). The storm center has shifted ~100 km to the east and the sea level pressure has dropped dramatically relative to the control run. At 1200 UTC 19 December, the storm central pressures in the MM5 (control) and LDA runs were 981 and 972 hPa, respectively, whereas the analyzed value was 972 hPa. The GFS (Eta) Model 24-h forecast valid at 1200 UTC 19 December showed 984 (983)-hPa central pressure, corresponding to 12 (11)-hPa error. During the first 9 h of the forecast, the pressure in the LDA run stays within 2 hPa of the control run and analyzed values (Fig. 13). Thereafter, pressure starts to drop steeply. The rate of deepening is similar to that analyzed, but the LDA run deepens the storm ~3 h earlier than analyzed. This is in contrast with the shorter run discussed in section 3a in which the LDA run started to deepen the storm 3 h later than analyzed. The pressure in the LDA run remains slightly lower than that analyzed after the storm starts to weaken (hours 27–36 in Fig. 13).

It is noteworthy that the 36-h control run predicted the deepening of the storm better than the shorter 24-h control run (Figs. 3 and 13). The separation occurred between 0900 and 1200 UTC 19 December when the 24-h run already started to weaken the storm, whereas the 36-h run continued to deepen the storm. The two storm centers were farther apart in the 36-h run and the northwest center was deeper. In addition, the 36-h run kept the two centers apart longer before merging them than did the 24-h run. This may have impacted the timing and rate of deepening of the storm in the control run.
Data from a QuikSCAT overpass at 0340 UTC 19 December were available to validate the modeled storm location and wind field. The western center has a wind speed minimum at 44.5°N, 139.4°W although the wind vectors suggest more northwesterly position, maybe 45.7°N, 140.5°W (Fig. 14c). As discussed in section 2c, the wind direction has large uncertainty at low wind speeds. In either case, both the control and LDA forecasts are quite similar and place the western center too far to the north and east, with LDA run having a slightly better location (Figs. 14a,b). If the 40-min time difference between the satellite overpass and forecast time is ignored, the location errors for control and LDA forecasts are ~325 and 280 km, respectively. It appears from the QuikSCAT data that there are very strong northerly winds between the two low centers (although some vectors were rain-flagged), suggesting that the western center has weakened more than both the model results show.

The eastern center has a wind speed minimum at 44.4°N, 135.2°W, although the wind vectors suggest a position ~100 km to the northwest. The control run places the eastern center ~1.5° too far to the south and ~4° to the east—which corresponds to a location error of ~290 km. The LDA forecast locates the center ~1° to the south and ~5° to the east of the observed—equivalent to a ~370-km location error. However, the pressure of the eastern center is 986 (977) hPa for the control (LDA) run (Fig. 13). This results in an enhanced wind field in the LDA run that matches better with observations, specifically on the eastern side of the low center. This also supports the improved surface pressure field in the LDA run.

The location of the storm from 36-h model forecasts (valid at 0000 UTC 20 December) was compared with data from the QuikSCAT overpass at 0310 UTC 20 December, discussed in section 3a. The QuikSCAT data place the storm center at 44.3°N, 131.6°W, whereas both the control and LDA forecasts put the storm center too far to the north and east. The storm center coordinates from the control run were 46.0°N, 130.6°W and from the LDA run were 47.8°N, 130.0°W. The more northeasterly location of the LDA forecast may have happened because the strongest advection of high-$T_v$ air occurred over the eastern and northern part of the storm center, thus contributing to falling pressure on that side of the storm.

c. 48-h model run

An attempt was made to start the model run at 0000 UTC 18 December 2002 and similarly assimilate lightning data for the first 8 h of the model run. Lightning over the storm started approximately at 0000 UTC, but only a few flashes were detected during the first 6 h (Fig. 2b). More lightning was observed during the last couple of hours of the assimilation period, although these observations were far south (~700–1000 km) of the storm center. The control forecast valid at 1200 UTC 19 December shows a storm central pressure of 981 hPa, whereas the LDA run shows 983 hPa. The analyzed value was 972 hPa (Fig. 2g). The storm center in the LDA run is farther east than in the control run and observations. The LDA run produces a larger and deeper secondary center to the west than the control run. Time series of the central pressure show that the LDA run tracks closer to (or the same as) the observations than the control run does during the first 21 h, although the differences are relatively small (Fig. 15). The separation occurs at forecast hour 30, at 0600 UTC 19 December, when the two low pressure centers merged in the control run. The LDA run kept the two centers separate, the storm center larger, and the central pressure higher than the control run or shown by the observations. The reason for the poorer performance of this LDA run may be the result of the relatively low lightning activity in the vicinity of the storm during the period of data assimilation (Fig. 2b).
d. Sensitivity tests

Errors in lightning and rainfall measurements impact the latent heating rates that are assimilated into the model. Errors in lightning measurements have a direct effect on the rainfall and latent heating rates that are assimilated (Fig. 1; x-error bars). On the other hand, the standard deviation in convective rainfall rates is quite large (Fig. 1; y-error bars) and it is important to know the impact of these uncertainties on modeling results. Errors in quantified lightning rates may result from uncertainties in the detection efficiency (DE) model, terminator effects that influence the sferics propagation between the day and night sides of the earth and power or network outages that disable a sensor location.

The sensitivity of the model simulations to rainfall rates was investigated by using the upper and lower boundaries for convective rainfall rate shown in Fig. 1. These curves correspond to ± one standard deviation in rainfall rates. The results from these two runs (hereafter referred as +stdev and −stdev) were compared to the “standard” LDA model run using the middle curve in Fig. 1.

The difference between the standard LDA run and the −stdev run for the 24-h forecast is negligible (Fig. 16a). The difference in sea level pressure field is less than 1 hPa over the whole domain and the difference in rainfall rates less than 3 mm (per 3 h).

When the +stdev run was compared with the standard LDA run, the differences were slightly larger but still minor. The differences in the rainfall field were generally small, except near the storm center, at 45°N, 132°W (Fig. 16b). It should be noted that the higher rainfall rates in the standard LDA run (yellow–red colors) are accompanied by higher rain rates in +stdev run (blue–white colors), indicating a shift in rainfall pattern rather than just a difference in rain rates. The difference in sea level pressure field is generally less than 1 hPa, except near 46.5°N, 132°W, where it reaches 5 hPa. It should be emphasized that this difference is not the difference in storm central pressure. That difference is located just
west of the storm center, where the pressure gradient was very high and the isobars tight (Fig. 12b). Thus, even a slight variation in storm location is emphasized in Fig. 16b. The difference in storm central pressure was 2 hPa.

Figure 17 shows that the difference in central pressure during the 36-h forecast period remains below 2 hPa in all three LDA runs. The $+\text{stdev}$ run has slightly lower pressure than the standard LDA run. This is understandable because the increased latent heating forcing results in increased $T_v$ values over the storm center and lower central pressure. Interestingly, the $-\text{stdev}$ run did not raise the storm central pressure above the standard LDA run; rather, it kept it the same or 1 hPa lower. The reason for the model’s insensitivity to increased rain

---

**FIG. 12.** The 24-h MM5 forecast of sea level pressure (contours; hPa) and 3-h accumulated rainfall (colors; mm) valid at 1200 UTC 19 Dec 2002: (a) Control, (b) LDA run, and (c) values of (b) − (a).
rates may be in the assimilation method. The latent heating scaling coefficient is limited to 3 (section 2b), resulting in the same heating rates for all the rainfall rates that reach the threshold.

The x-error bars in Fig. 1 show a hypothetical ±50% error in lightning rates. The corresponding two curves show that a ±50% error in lightning rates has very small impact on the assimilated rainfall rates. Thus, it is expected that an error of this magnitude has much less impact than a ±1σ error in rainfall rates and does not change the model results significantly.

The analysis shows that the model is relatively insensitive to errors in the lightning–rainfall relationship and is also insensitive to assimilated lightning rates. This result likely reflects the fact that even ±50% variations in lightning rates result in small variations in the rainfall and assimilated latent heating rates because of the logarithmic nature of the lightning–rainfall relationship (Fig. 1). Moreover, triggering convection in the grid points where the model is originally dry is presumably more significant than small errors in the estimation of latent heat. These results are encouraging for lightning data assimilation because they relax the requirements for accuracy in quantifying lightning rates (DE model) and in deriving the lightning–rainfall relationship.

To test the sensitivity of the length of the LDA period on the forecasts, the model was run with only 3 h of LDA as

![Figure 13](image13.png)

**FIG. 13.** As in Fig. 3, but the forecast was initiated at 1200 UTC 18 Dec 2002 and run for 36 h.

![Figure 14](image14.png)

**FIG. 14.** As in Fig 5, but for 15-h MM5 forecasts; forecasts are valid at 0300 UTC 19 Dec 2002. QuikSCAT wind field is for 0340 UTC 19 Dec 2002.
opposed to 8 h. The 24-h and 36-h LDA forecasts deepened the storm to 979 and 977 hPa, whereas the standard LDA forecasts were 975 and 971 hPa, respectively.

4. Summary and conclusions

Midlatitude low pressure systems that propagate eastward over the North Pacific Ocean can bring hazardous weather conditions to the West Coast of North America. The paucity of in situ observations over the Pacific Ocean can lead to inadequate initial conditions in NWP models and result in poor forecasts of the storms. VLF sferics observations from PacNet show that winter storms over the North Pacific Ocean are often associated with significant lightning activity. These lightning observations offer important additional information about convective activity and diabatic heating over the data-sparse ocean. In this paper, the potential of lightning data assimilation to improve NWP forecasts over the North Pacific Ocean is investigated using, for the first time, a continuous, calibrated lightning data stream from the Pacific Lightning Detection Network/Long-Range Lightning Detection Network (Pessi et al. 2009).

The lightning data assimilation method described in this paper utilizes an empirical lightning–rainfall relationship derived previously from lightning and satellite data (Pessi and Businger 2009). The method adjusts the vertical latent heating profiles in MM5’s Kain–Fritsch convective parameterization scheme according to rainfall rates estimated from lightning observations. The method is designed to be employed operationally, allowing 8 h of assimilation at the beginning of the model run. The assimilation method can be easily ported to another numerical model if the (KF) convective parameterization module is available for the model. PacNet data are currently being assimilated operationally.
into the Weather Research and Forecasting (WRF) model at the Mauna Kea Weather Center using an improved method based on this paper.

To illustrate the promise of the approach, lightning data from a notable storm that occurred over the Northeast Pacific Ocean in late December 2002 were assimilated into the MM5. The focus of the case study presented in this paper is the impact of LDA on the synoptic- and subsynoptic-scale pressure distribution and wind field forecast surrounding the storm center, rather than on changes in mesoscale rain patterns or quantitative precipitation forecasts (QPFs). Over the open ocean, accurate wind forecasts are critical for maritime, coastal, and aviation interests. Forecast wind fields in this study were verified using QuikSCAT, ship, and buoy data.

The control forecast, without lightning data, was unable to predict the rapid deepening of the storm. The MM5 forecast errors for the time of observed deepest storm central pressure (valid 1200 UTC 19 December) were 10 and 9 hPa in the 12- and 24-h forecasts, respectively. In contrast, the lightning data assimilation run predicted the central pressure within 5 hPa of that analyzed in the 12-h forecast and correctly predicted the central pressure in the 24-h forecast.

A detailed analysis of the modeled storm dynamics revealed that enhanced latent heating increased the temperature gradient across the cold front, resulting in an increase in along-front winds, consistent with thermal wind considerations. The enhanced flow increased the advection of warm air over the storm center, resulting hydrostatically in a drop in sea level pressure.

It is interesting to note how the rate and timing of the deepening of the storm using LDA matches the analyzed values. In Fig. 3, the deepening of the storm occurs simultaneously with the LDA period, although there is a ∼3-h lag in deepening compared to analyses. In contrast, Fig. 13 shows that the LDA run started to deepen the storm after the assimilation period had ended. However, in both cases the rate of the deepening was relatively well simulated.

Sensitivity tests showed that the model is relatively insensitive to errors in the lightning–rainfall relationship and very insensitive to assimilated lightning rates. The logarithmic distribution of the lightning–rainfall relationship implies that even ±50% variations in lightning rates result in very small variations in the rainfall and assimilated latent heating rates. These results are consistent with those of Manobianco et al. (1994) and Chang et al. (2001), who also found that assimilating latent heating in the correct location is more important than accurate estimation of the precipitation intensity.

The LDA along the cold front resulted in long-term (>36 h) impact on the synoptic-scale storm development. The results of this study suggest that assimilation of a calibrated, long-range lightning data stream in an operational forecasting system may improve forecasts of cyclogenesis over the open ocean. However, to confirm this initial promise, a rigorous comparison of model performance with and without lightning data in an operational setting needs to be undertaken. This test is currently under way over the North Pacific Ocean.

Acknowledgments. We thank Tiziana Cherubini for her assistance in the modeling experiments. We are grateful to Nick Demetriades for bringing this storm case to our attention. The QuikSCAT data were obtained from the Physical Oceanography Distributed Active Archive Center (PO.DAAC) at the NASA Jet Propulsion Laboratory, Pasadena, California (see http://podaac.jpl.nasa.gov). This work is supported by the Office of Naval Research under Grants N000140510551 and N000140810450.

REFERENCES


